Journal of Real-Time Image Processing
https://doi.org/10.1007/s11554-019-00917-3

SPECIAL ISSUE PAPER

Coverless real-time image information hiding based on image block matching and dense convolutional network

Yuanjing Luo[1] · Jiaohua Qin[1] · Xuyu Xiang[1,2] · Yun Tan[1] · Qiang Liu[1] · Lingyun Xiang[3]

Received: 25 July 2019 / Accepted: 12 September 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Information security has become a key issue of public concern recently. In order to radically resist the decryption and analysis in the field of image information hiding and significantly improve the security of the secret information, a novel coverless information hiding approach based on deep learning is proposed in this paper. Deep learning can select the appropriate carrier according to requirements to achieve real-time image data hiding and the high-level semantic features extracted by CNN are more accurate than the low-level features. This method does not need to employ the designated image for embedding the secret data but transfer a set of real-time stego-images which share one or several visually similar blocks with the given secret image. In this approach, a group of real-time images searched online are segmented according to specific requirements. Then, the DenseNet is used to extract the high-level semantic features of each similar block. At the same time, a robust hash sequence with feature sequence, DC and location is generated by DCT. The inverted index structure based on the hash sequence is constructed to attain real-time image matching efficiently. At the sending end, the stego-images are matched and sent through feature retrieval. At the receiving end, the secret image can be recovered by extracting similar blocks through the received stego-images and stitching the image blocks according to the location information. Experimental results demonstrate that the proposed method without any modification traces provides better robustness and has higher retrieval accuracy and capacity when compared with some existing coverless image information hiding.

Keywords Coverless information hiding · Data hiding · Deep learning · DCT · DenseNet · Real-time image processing

1 Introduction

Due to the wide use of multimedia data (digital image, audio, and video) in our electronic world, the communication of secret information in a digital way is urgently required [1]. Information hiding is the art of communicating secret information in a hidden manner [2]. Generally, it hides secret information imperceptibly into an appropriate host medium such as digital image, audio, and video files so that the presence of hidden information cannot be diagnosed [3]. The traditional image information hiding designates a cover image and embeds the secret information into the carrier by making the slight modification to the carrier data (digital image), and the secret data to be hidden are replaced at pixels of the carrier image [4]. However, these modification traces caused by the embedding will be left in the cover image, which allows the hidden information to be detected [5]. To radically resist the detection technology and improve the robustness [6] of image information hiding, Bilal et al. proposed "Zero-steganography" in 2013 [7]. To improve security, Zhou et al. proposed the new concept of "coverless" in May 2014 [8]. It does not need to designate and modify a cover image to hide the secret information. Instead, the hiding process is implemented by finding an image or text that already contains the secret information. As we know, any image contains a lot of information. It is possible to map some relationships between these feature information and secret information to be hidden with a proper feature description [9]. The common

✉ Jiaohua Qin
qinjiaohua@163.com

✉ Xuyu Xiang
xyuxiang@163.com

[1] College of Computer Science and Information Technology, Central South University of Forestry and Technology, Changsha, China

[2] College of Communication and Information Sciences, University of Alabama, Tuscaloosa, USA

[3] School of Computer and Science Engineering, Changsha University of Science and Technology, Changsha, China

Published online: 23 September 2019

method is to build mapping relationships between the hash sequences and the secret messages [10]. Consequently, if we hide n-bit secret information into an image, it needs 2^n natural images to transmit secret data. The number of images increases exponentially with the length of the secret data, which makes those approaches need to be improved.

For a given secret image, it is possible that some visually similar blocks may exist between the secret image and other unrelated images. If we can find a set of similar blocks of the secret image in a natural image data set, we can use these unrelated images as stego-images to represent and hide the secret image. Therefore, Zhou et al. proposed a novel coverless steganography approach for transmitting the secret image using a set of proper similar blocks of a given secret image as stego-images [11], which are retrieved from a natural image database in the meantime. At the receiver end, the similar regions can be cropped from the stego-images and spliced together to recover the secret image approximately. However, this method has the following disadvantages:

1. The feature extracted from the image block by this method is the gray histogram, which leads to the less robustness to Gaussian noise. When the restored image is attacked by Gaussian noise, it is not visually similar to the original secret image. As a result, the security of this method is greatly reduced.
2. When constructing the index structure, the indexes of the two layers are the sequence codes constructed by the average pixel value and the gray value, respectively. In the first classification, the image blocks with the same location information cannot meet the requirements of feature matching to be the similar blocks. Therefore, the retrieval accuracy still has room for improvement.

In this paper, we present a coverless real-time image information hiding based on image block matching and Dense Convolutional Network. In our approach, we have made the following optimization for Zhou's work aiming at the above two defects to improve the accuracy and robustness:

1. Instead of building a local image database, we use real-time search to cut and match the image blocks on the Network.
2. Deep Learning is one of the research fields of Machine Learning [12]. It can achieve end-to-end supervised learning and unsupervised learning by establishing a hierarchical artificial neural networks. Convolutional neural network (CNN) usually extracts high-level semantic features which are generated by convolution, pooling, and other operations; we use the DenseNet to extract the repeated high-level semantic features of each block instead of the gray value. Such features can simulate human perception to some extent, and thus achieve higher accuracy in retrieval or recognition than low-level features such as SIFT and SURF [13].
3. We adopt the supervised learning of deep learning to select the appropriate carrier from image set according to our content requirements, which reduces the risk of a random selection of carrier, greatly saves the time of image matching and can achieve the real-time image matching.
4. We use DCT to generate the sequence code from the image, which is more robust than the sequence code generated by simple z-shape arrangement.

The rest of this paper is organized as follows: Sect. 2 introduces the related works, including the definition of the stego-image, the DenseNet, and DCT. In Sect. 3, we will introduce coverless real-time image information hiding based on image block matching and Dense Convolutional Network. We will evaluate the performance of this method by comparing it with other methods in terms of robustness, accuracy, efficiency and capacity in Sect. 4. Finally, a summarization of this method and the plan for the next step is given in Sect. 5.

2 Related works

2.1 Stego-images

The image itself already contains a lot of feature information, such as pixel brightness value, color, texture, edge, contour and advanced semantics. For a given image, there may be some visually similar parts between the image and another image in the natural image data set, as shown in Fig. 1. The given image on the left and the natural image on the right can be viewed as sharing the same image block. The original image on the right, which includes the similar block is called the stego-image. If we can find a set of similar blocks in the natural image data set, we can use these stego-images containing similar blocks to represent the secret image. At

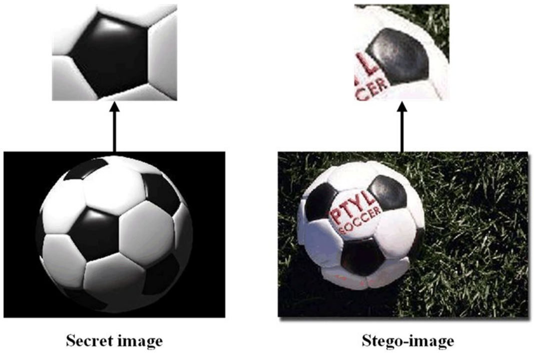


Fig. 1 Secret image and stego-image

the receiving end, these image blocks can be clipped from the stego-image and stitched together to restore the same image as the secret image visually. Note that because these image blocks are retrieved from the natural image data set, information hiding can be achieved without modification. In other words, only natural images without any information are transmitted, which can achieve the effect of secret image transmission.

2.2 DenseNet

Convolutional neural network is a type of deep learning model for processing data, which is inspired by the organization of animal visual cortex and designed to automatically and adaptively learn spatial hierarchies of features from various levels [14]. The convolutional neural network is a very popular network framework in the field of deep learning, especially in the field of computer vision. Since ResNet was proposed and became a milestone in the convolutional neural network, the various networks of Resnet emerge in endlessly and the performance have been improved to some extent. As the latest Network, DenseNet (Dense Convolutional Network), its structure is not complicated, but very effective [15]. It can be said that DenseNet absorbed the essential part of ResNet and improved performance. DenseNet is a convolutional neural network with dense connections. In this network, there is a direct connection between any two layers. The input of each layer of the network is the union of the output of all previous layers, and the feature graph learned by this layer will be directly transmitted to all subsequent layers as input. The system architecture diagram of DenseNet is shown in Fig. 2. It can be found the reference of this network alleviates the problem of gradient disappearance, enhances feature propagation, encourages feature reuse, and greatly reduces the number of parameters.

2.3 Discrete cosine transform

Discrete cosine transform is a typical transformation method with eight forms. In particular, its second type is often used in signal processing and image processing for lossy data compression of signals and images (including still images and moving images) [16]. For example, the image is encoded after the 8×8 blocks DCT transformation in the image encoding standard JPEG. Assuming the length of signal $f(x)$ is a and the second form of DCT transformation is expressed as

$$C(u) = \alpha_u \sum_{x=0}^{a-1} f(x) \cos \frac{\pi(2x+1)u}{2a}, \tag{1}$$

where $C(u)$ is the transform coefficient, u is the frequency and α_u can be considered as a compensation coefficient, which can be formulated as

$$\alpha_u = \begin{cases} \sqrt{1/a}, & u = 0 \\ \sqrt{2/a}, & 1 \leq u \leq a-1 \end{cases} \tag{2}$$

Since the image is a two-dimensional matrix, DCT transformation needs to be extended to a two-dimensional form. Thus, the DCT coefficients of an image $f(x, y)$ with a size of $a \times b$ can be computed according to the following equation. Therefore, the following transformation is required.

$$C(u,v) = \alpha_u \alpha_v \sum_{x=0}^{a-1} \sum_{x=0}^{b-1} f(x) \cos \frac{\pi(2x+1)u}{2a} \times \cos \frac{\pi(2y+1)v}{2b}, \tag{3}$$

where $C(u, v)$ represents the coefficient after image transformation, u and v represent horizontal and vertical frequency, respectively. α_u and α_v are defined as

$$\alpha_u = \begin{cases} \sqrt{1/a}, & u = 0 \\ \sqrt{2/a}, & 1 \leq u \leq a-1 \end{cases} \tag{4}$$

$$\alpha_v = \begin{cases} \sqrt{1/b}, & u = 0 \\ \sqrt{2/b}, & 1 \leq u \leq b-1 \end{cases} \tag{5}$$

DC coefficients can be calculated by Eq. (6)

$$\mathrm{DC}(u,v) = \frac{1}{\sqrt{ab}} \sum_{x=0}^{a-1} \sum_{x=0}^{b-1} f(x). \tag{6}$$

According to the above, DC is calculated with each pixel of the image. The change of DC coefficient will significantly reduce the image quality, resulting in poor image insensitivity. Because the coverless image steganography hides secret

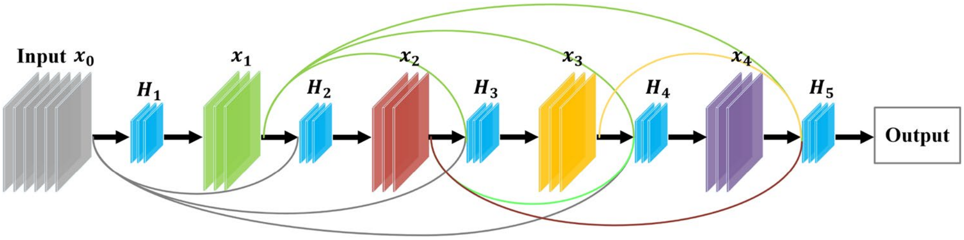


Fig. 2 The system architecture diagram of DenseNet

information by establishing mapping rules. Therefore, its modification does not affect the carrier image.

3 The proposed information hiding

In this section, we will demonstrate the process of hiding and extracting secret images. Figure 4 shows a diagram of our approach. With more and more attention paid to timeliness, real-time image processing has become a hot issue in the field of image processing [17]. We choose real-time image search on the network for processing, achieving an ideal effect in terms of efficiency. In this framework, a large number of images on the network are searched and downloaded in real-time [18]. Then, we divided these real-time images into several non-overlapping blocks, extracted the high-level semantic features of each block with DenseNet for feature matching [19]. Based on this, we used DCT to generate a robust hash sequence through the DC coefficients between adjacent image blocks and constructed the inverted index structure. To hide the real-time image, we first segment the image into appropriate blocks. Then, for each image block, we use the inverted index to retrieve a natural image containing a similar block. Therefore, we get a lot of stego-images, which are transmitted to the receiver. At the receiving end, we also use location information to extract these image blocks from the stego-images. Finally, we can splice these image blocks together according to the location information to restore the secret image. In general, the main parts of this method include the feature extraction, the generation of hash sequence code, the construction of the inverted index and the transmission of secret information.

3.1 Feature extraction

Convolutional neural network (CNN) usually extracts high-level semantic features which are generated by convolution, pooling, and other operations, and can simulate human perception to a certain extent [20]. To ensure the real-time performance of the algorithm, deep learning network is fully trained in advance [21]. In this paper, we used ImageNet data set to train DenseNet121 network. Considering the security risks caused by the random selection of carrier image and the requirements of carrier image in a specific type, we adopt the supervised learning to select the appropriate carrier from image set according to our content requirements. In this research, we adopted the DenseNet-based feature retrieval scheme and used supervised learning to retrieve corresponding types of images (one or more classes) as the transmission carrier, compared with the low-level features such as SIFT and SURF, the scheme achieves a higher accuracy rate in the field of retrieval. The feature retrieval process is as follows:

1. DenseNet convolutional neural network is used to extract features from the image set. For a given image x, features $F_{(x)} = \{f_1, f_2, \dots, f_a\}$ are extracted from the global average pooling layer, where its dimension is $1 \times a$.The extracted features F'_x are obtained after Min-Max normalized.

$$F'_x = \frac{f_x - f_{\min}}{f_{\max} - f_{\min}}, \quad n = 1, 2, \dots, a, \tag{7}$$

where f_x is the feature of one dimension,$f_{\min}$ is the min feature of all dimensions and $f_{\max}$ is the max feature of all dimensions.

2. Similarly, the features F'_y from image set $Y = \{y_1, y_2, \dots, y_n\}$ can be obtained after normalized according to the Eq. (7)

$$F'_y = \left\{ F'_{y_1}, F'_{y_2}, \dots, F'_{y_n} \right\}. \tag{8}$$

3. The similarity between F'_x and F'_y is measured by Euclidean distance.

$$\begin{aligned} \mathrm{Sim}(F'_x, F'_{y_m}) &= \mathrm{Distance}(F'_x, F'_{y_m}) \\ &= \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}, \quad 1 \le m \le n, \end{aligned} \tag{9}$$

where $\mathrm{Sim}(F'_x, F'_{y_m})$ is the Euclidean distance between x and mth images in the image set.

4. The corresponding image is retrieved according to the similarity.

$$P_m = \mathrm{Sort}(F'_x, F'_{y_m}), \tag{10}$$

where P_m is the image to be matched for the image to be retrieved.

3.2 Generation of hash sequence code based on DCT

Hash feature sequences extracted from images have a fixed length [22]. An image can be represented by feature sequence due to the mapping relationship between feature sequence and secret information segment. We usually extract feature sequences from images according to feature extraction methods. However, sequence codes are very weak and sensitive to geometric attacks (rotation, translation) and processing (adding noise, compression, and brightness variations). To ensure the secret information can be successfully extracted at the receiver, the image feature sequence must be robust [23].

DCT coefficients include DC coefficients and various ac coefficients. Most of the energy and content is contained in the low-frequency coefficients. For high-frequency coefficients, they can represent the details of the image and obtain higher image quality compared with the low-frequency coefficient. In our scheme, we extract feature sequences in the DCT domain, and the specific steps are as follows:

1. We first resize all images $M \times N$, then divide them into $m \times n$ blocks, finally obtain the image block B_i.

$$B_i = \left\{B_1, B_2, \dots, B_{mn}\right\}, \quad 0 \le j \le mn. \tag{11}$$

2. For each B_j, we convert it to YUV and divide the Y channel into 16 subblocks b_j which can be obtained by raster scanning order.

$$b_i = \left\{b_1, b_2, \dots, b_{16}\right\}, \quad 0 \le j \le 16. \tag{12}$$

3. For each b_j, we apply DCT transform to it and get the adjacent subblock D_j.

$$D_j = C_s(u, v) = DCT(b_j), \quad 0 \le j \le 16. \tag{13}$$

4. We get each bit of the sequence code of B_i according to $D_s i$.

$$f_{b_j} = \begin{cases} 1, & \text{if } D_j > D_{j+1} \\ 0, & \text{if } D_j \le D_{j+1} \end{cases}, \quad 0 \le j \le 16. \tag{14}$$

5. Repeat the above steps to obtain the hash sequence code of all similar blocks in the image.

Algorithm 1 DCT binarization algorithm.

```
Input: An image block $B_i$.
Output:The hash sequence code.
Convert $B_i$ to YUV.
Divide the Y channel into 16 subblocks $\{b_1,b_2,...b_{16}\}$
for $j$ =1,...,16 do
Calculate the adjacent subblock $D_j$ according to Eq(13).
if $D_j$>$D_{j+1}$
    $f_{b_j}$=1
else
    $f_{b_j}$=0
end if
```

Algorithm 2 Generation of hash sequence code.

```
Input: An image with $M \times N$ size.
Output: The hash sequence code of all similar blocks.
Divide the image into $m \times n$ blocks $\{b_1,b_2,...b_{16}\}$.
for $j$ =1,..., $m \times n$ do
Calculate the hash sequence of the block according to Algorithm 1.
Get the hash sequence code of all similar blocks.
```

The DCT binarization algorithm is used to generate the hash sequence code which can be determined by Algorithm 2. Note that to consider the capacity and robustness, we extract the first eight bits as the feature sequence of each B_i. The hash sequence of all similar blocks is generated according to Algorithm 2.

3.3 Construction of inverted index

Although we adopted the method of supervised learning and selected the appropriate carrier to retrieve the image library according to our own content requirements, we still need to speed up the matching of secret information with images further, the index structure is created. According to these steps above, each similar block can be represented by eight-bit hash sequence code, and image blocks with the same sequence code have similarity in features. As shown in Fig. 3, we first index all the images from the database according to their hash sequence. Then, we build an inverted index structure which contains all possible 8 bits hash sequences as entries. Under each entry is a set of paths to the image block, which can be used to find the image block. Note that there should be at least one image under each list of sequence codes to ensure that for any combination of sequence codes, the corresponding image can be found in the index structure.

3.4 Information hiding and extraction

As shown in Fig. 4, information transmission includes the following four steps: (1) it is similar to the processing process of the database image. For a secret image that needs to be hidden, we divide it into blocks. (2) To improve the retrieval accuracy and security of this method, we adopted

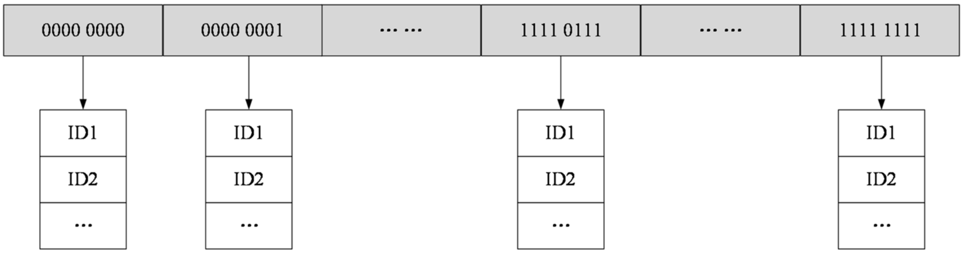


Fig. 3 The inverted index

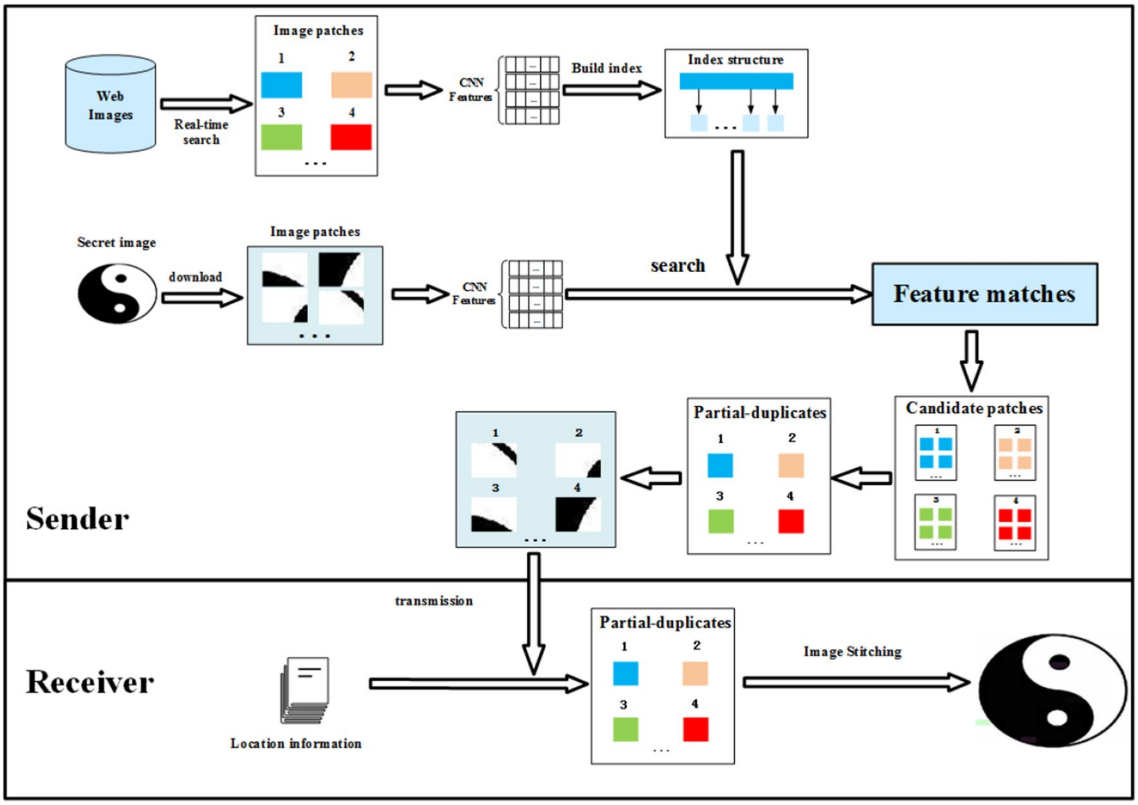


Fig. 4 Coverless real-time image information hiding based on image block matching and dense convolutional network

higher-level semantic features which are more accurate than SURF. (3) Through the above steps, we can use the index file to obtain many similar blocks from the image database. At the receiving end, we use the location information to obtain similar blocks from the stego-image. Next, a blank area of the same size as the secret image is given. Finally, these image blocks are placed into the blank area according to the position to generate an image similar to the secret image. Through the above steps, we can use the inverted index to get a number of image blocks from the image set. For better restoration effect, we use Euclidean distance to calculate the similarity between the candidate image blocks and secret image block instead of randomly selecting them.

4 Experimental results and analysis

We conduct our experiments on a standard computer with GPU due to the introduction of deep learning. Before transmitting the secret information, each secret image is resized into 640 × 480 and the secret image is cut into 100 blocks for the results of our collection. The feature of the image block is extracted through the trained DenseNet, the dimension a is 1024. Figure 5 shows the recovered secret image through our method. As we have known, coverless information hiding aims at the secret hidden information which is necessary to pay attention to the following aspects: first, the feature extracted cannot be single. Otherwise, the capacity and efficiency of data transmission will be insufficient. Second, to send information quickly and accurately, the sender usually needs to prepare the image set to retrieve which is formed by a large number of natural images in advance [24], and these natural images come from a wide range of sources, which cannot correctly meet the ideal situation. Third, the image is vulnerable to attacks in the transmission process, resulting in the disclosure of secret information. It must be able to resist steganalysis and attackers. To sum up, in this section, we will evaluate our method from the above four aspects: capacity, efficiency, accuracy, and safety.

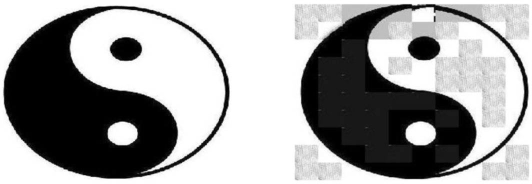


Fig. 5 The original secret image (left) and the restored image (right) (100 blocks)

4.1 Capacity

Coverless information hiding has many advantages, but it still has the problem of capacity. Although the image itself already contains a lot of feature information, such as pixel brightness value, color, texture, edge, contour, and semantics. The common information hiding method is to construct a hash sequence code using SIFT features to map the secret data [25]. However, the capacity is limited by the image hash length [26]. Therefore, the capacity of the uncovering method is obviously smaller than that of the traditional image information hiding. In our method, a 15-bit sequence code can be generated, but we only use the first 8 bits. In general, an image block can carry up to 15 bits of information. As shown in Table 1, the capacity of the whole image is determined by the number of image blocks.

There are four typical coverless steganography methods: Zhou et al. proposed three coverless steganography methods [3, 8, 11], Yuan et al. proposed a coverless image steganography based on SIFT and BOF [27]. In our experiment, we cut the image into 100 blocks and extracted the first 8 bits of sequence code to hide information and we compare our embedded capacity with the four existing coverless information hiding. As shown in Table 2, it is not difficult to find that our method is superior to the existing information hiding method in terms of capacity.

Table 2 The capacity comparison

Method	Capacity (bit)
Ours	**800**
Zhou [3]	8
Zhou [8]	8
Zhou [11]	384
Yuan [27]	16

Bold is to highlight the performance of this method

4.2 Efficiency

In the field of data hiding, how to deal with a large amount of multimedia data quickly has become a new challenge. More and more image information hiding focuses on the improvement of efficiency. Our method mainly spends time in the following aspects: sequence code generation and feature matching [28]. Compared with Zhou [11], since we generate sequence codes from images through DCT changes instead of directly extracting pixel values for matching, it increases the robustness and time consumption at the same time. We tested the time consumption of secret image SI1 and SI2. As shown in Table 3, time consuming is the time to extract each feature sequence. It can be seen that our method takes a little more time in feature matching of blocks. However, in the whole process of image information hiding, the gap can be ignored. On the contrary, in general, our method has two advantages in terms of efficiency:

Table 1 Embedding capacity of our approach

Number of blocks	Capacity (bit)
40 × 30	9600–180,000
20 × 15	2400–4500
8 × 6	384–720
4 × 3	96–180

1. Combined with deep learning, feature selection based on DenseNet can filter out most unnecessary images from the natural data set in advance, saving the retrieval time.
2. The construction of reverse index can quickly match the required image blocks, saving most time.

Because real-time image processing is becoming a hot topic in this field, real-time frame processing rate has also become an important evaluation index. For a real-time image, our method can extract feature sequences in blocks in a relatively short time. At a word, as long as we carry out feature processing on the required image, the time to extract each feature sequence is 0.044, close to 0. The information extracted by the receiver can achieve good results.

4.3 Accuracy

The visual similarity between the restored image and the original secret image can be used to evaluate the accuracy of this method. The SSIM (Structure Similarity) index is an image quality evaluation method that the evaluation result seems more consistent with the visual perception of people. In this section, we use SSIM to evaluate the accuracy of our method. The SSIM index is a full-reference image quality measuring method that it can compare the illuminance, contrast, and structure of two images to calculate their similarity. The SSIM index is in the range of [− 1, 1], and when it equals to one, the accuracy is the best. In our experiment,

Table 3 Time consuming of different method

Method	Time consuming of SI1 (s)	Time consuming of SI2 (s)
Ours	**0.0449**	**0.0441**
Zhou [11]	0.0217	0.0227

Bold is to highlight the performance of this method

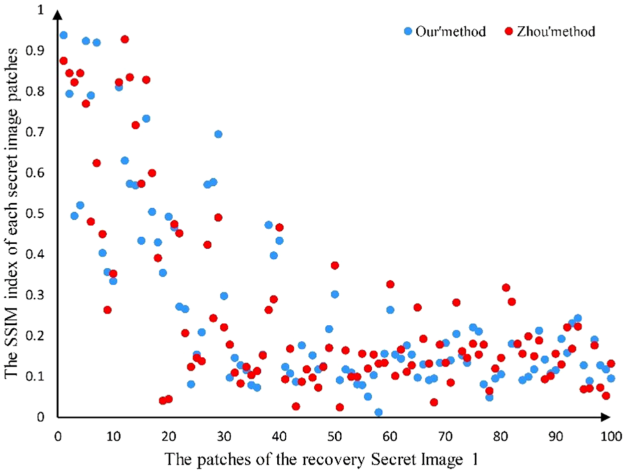


Fig. 6 The SSIM index of image patches of SI1

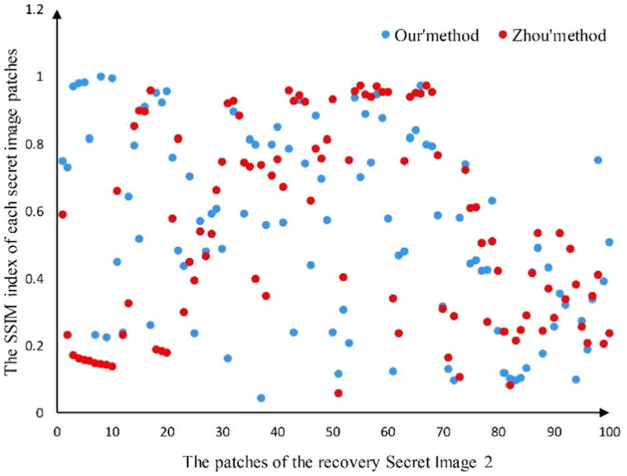


Fig. 7 The SSIM index of image patches of SI2

Table 4 The SSIM index of different methods

Method	SSIM of SI1 (%)	SSIM of SI2 (%)
Ours	**53.41**	**53.49**
Zhou [11]	25.51	35.05

Bold is to highlight the performance of this method

we selected two Secret Images downloaded from website and used SSIM to evaluate the similarity. The similarity of image patches is shown in Figs. 6 and 7, and the similarity of the images is shown in Table 4. In Figs. 6 and 7, the *x*-axis is the number of patches of the recovery image, and the *y*-axis is the SSIM index corresponding to each secret image patch and recovery image patches. The red dots represent Zhou [11], and the blue dots represent our method.

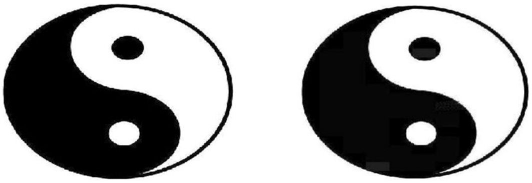


Fig. 8 The original secret image (left) and the restored image (right) (200 blocks)

From these experimental data, we can find that the accuracy of our method is slightly better than Zhou [11].

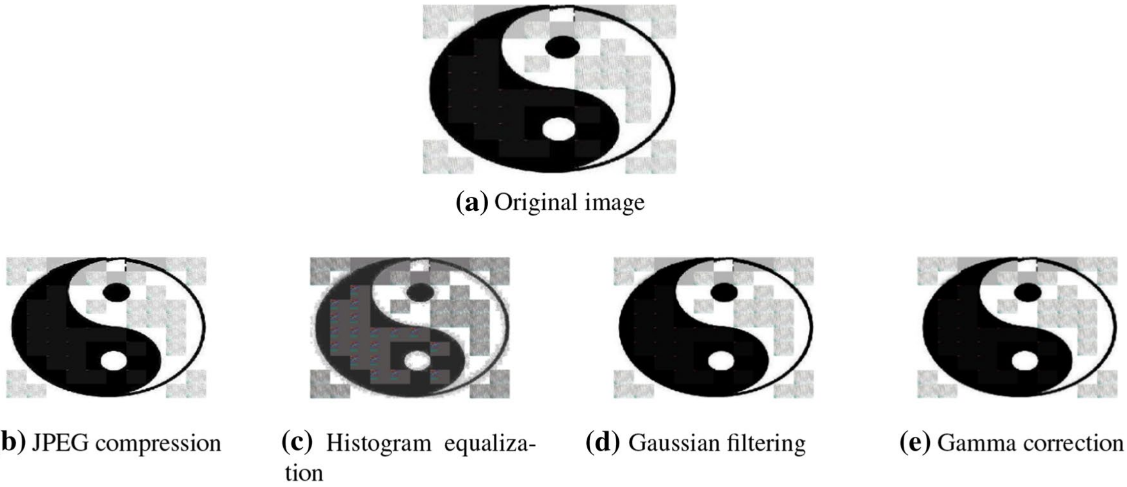
Meanwhile, the number of image blocks also affects the accuracy. When we divide the secret image into more image blocks, the more appropriate image block can be found to represent the secret image blocks so that the accuracy of the method will be better. As shown in Fig. 8, if we cut the secret image into 200 blocks for feature matching, the restored image is visually more similar to the original image.

However, if there are more image blocks, the capacity will become very low. Therefore, we need to strike a balance between the capacity and the accuracy and define the parameter according to our actual request. If we want to get a high accuracy, which means the restored image is the same as the secret image, we can divide the secret image into more image blocks. On the contrary, we can divide the secret image into fewer image blocks to hide more information.

4.4 Security

Anti-detection It consists secret information detection and extraction. Most detection methods determine the existence of secret information by analyzing the influence of embedded secret information on the statistical characteristics of images, such as the statistical anomaly of carrier data caused by information embedding. Most of the existing image information hiding embeds the secret information into the image by modifying the content or structure of the image. Therefore, detection tools can detect the existence of secret information through the modification traces left in the image. However, instead of modifying the content or structure of the image, we transmit a set of natural images that have nothing to do with the secret image and contain image block similar to the secret image. In short, our method embeds the secret image into the stego-image without any modification. Therefore, our method can resist the existing detection tools to a large extent.

Anti-attack (*robustness*) In the process of transmission, images will inevitably encounter a variety of content damage, such as rescaling, JPEG compression, Gaussian noise, histogram equalization and so on. The information extracted from the image must be able to resist these attacks. We use the success rate of secret data extraction to evaluate the


(a) Original image

(b) JPEG compression

(c) Histogram equalization

(d) Gaussian filtering

(e) Gamma correction

Fig. 9 The attack ways on the image

Table 5 The SSIM index of different attacks (%)

Attack	JPEG compression	Histogram equalization	Gaussian filtering	Gamma correction
Ours	**76 (SI1) 48 (SI2)**	**76 (SI1) 74 (SI2)**	**98 (SI1) 98 (SI2)**	**98 (SI1) 98 (SI2)**
Zhou [11]	66 (SI1) 39 (SI2)	79 (SI1) 68 (SI2)	98 (SI1) 98 (SI2)	93 (SI1) 91 (SI2)

Bold is to highlight the performance of this method

anti-attack ability of this method. Figure 9 shows several common image attacks.

1. JPEG compression with a fact of 10;
2. Histogram equalization;
3. Gaussian filtering with 3 × 3 window;
4. Gamma correction with 0.8. In this section, we choose two Secret Images and cut them into 100 patches to carry out experiments with our method and Zhou [11], respectively. We will use the SSIM index to verify the robustness of different ways of each image attack. The experimental results are given in Table 5, it can be found that our method is more robust than Zhou [11].

5 Conclusion

In this paper, we propose a coverless real-time image information hiding based on image block matching and Dense Convolutional Network. In our method, we extract higher-level semantic features based on supervised learning of deep learning, its perceptibility improves the accuracy of the method. Feature sequence is extracted from the image by DCT coefficient, which is a series of sequences with fixed length and strong robustness. At the same time, the inverted index structure is used to optimize the search efficiency. Compared with the existing methods, our method has higher accuracy and robustness, and can better protect the security of secret information.

Acknowledgements This work is supported by the National Natural Science Foundation of China (No. 61772561), the Key Research and Development Plan of Hunan Province (Nos. 2018NK2012, 2019SK2022), the Science Research Projects of Hunan Provincial Education Department (Nos. 18A174, 18C0262), the Postgraduate Research and Innovation Project of Hunan Province (No. CX20190625), the Postgraduate Science and Technology Innovation Foundation of Central South University of Forestry and Technology (No. CX20192016), and Science and Technology Innovation Platform and Talent Plan of Hunan Province (No. 2017TP1022).

References

1. Ma, W., Qin, J., Xiang, X., Tan, Y., Luo, Y., Xiong, N.N.: Adaptive median filtering algorithm based on divide and conquer and its application in CAPTCHA recognition. Comput. Mater. Contin. **58**(3), 665–677 (2019)
2. Qin, J., Sun, X., Xiang, X., Niu, C.: Principal feature selection and fusion method for image steganalysis. J. Electron. Imag. **18**(3), 1–14 (2009)
3. Zhou, Z., Sun, H., R.H., Chen, X., Sun, X.: Coverless image steganography without embedding. Cloud Computing and Security, pp. 123–132 (2016)
4. Ni, J., Ye, J., Yi, Y.: Deep learning hierarchical representations for image steganalysis. IEEE Trans. Inf. Forensics Secur. **12**(11), 2545–2557 (2017)
5. Gao, H.: Summary of research on key technologies of information hiding. Electron. World **9**, 146–148 (2016)

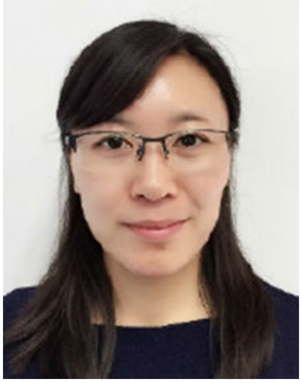
6. Tan, Y., Qin, J., Xiang, X., Ma, W., Pan, W., Xiong, N.N.: A robust watermarking scheme in YCbCr color space based on channel coding. IEEE Access. **7**(1), 25026–25036 (2019)
7. Bilal, M., Imtiaz, S, Abdul, W., Ghouzali, S.: Zero-steganography using DCT and spatial domain. In: 2013 ACS International Conference on Computer Systems and Applications (AICCSA) (2013)
8. Zhou, Z., Cao, Y., Sun, X.: Coverless information hiding based on bag-of-words model of image. J. Appl. Sci. Electron. Inf. Eng. **34**(5), 527–536 (2016)
9. Guo, Y., Li, C., Liu, Q.: R2N: a novel deep learning architecture for rain removal from single image. Comput. Mater. Contin. **58**(3), 829–843 (2019)
10. Zhou, Z., Jonathan, W., Sun, X.: Encoding multiple contextual clues for partial-duplicate image retrieval. Pattern Recognit. Lett. **15**(6), 1–9 (2017)
11. Zhou, Z., Mu, Y., Jonathan, W.: Coverless image steganography using partial-duplicate image retrieval. Soft Comput. **23**, 4927–4938 (2018)
12. Wang, J., Qin, J., Xiang, X., Tan, Y., Pan, N.: CAPTCHA recognition based on deep convolutional neural network. Math. Biosci. Eng. **16**(5), 5851–5861 (2019)
13. Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.: Speeded-up robust features (SURF). Comput. Vis. Image Underst. **110**(3), 346–359 (2018)
14. Xu, F., Zhang, X., Xin, Z., Yang, A.: Investigation on the Chinese text sentiment analysis based on convolutional neural networks in deep learning. Comput. Mater. Contin. **58**(3), 697–709 (2019)
15. Zhang, J., Lu, C., Li, X., Kim, H., Wang, J.: A full convolutional network based on DenseNet for remote sensing scene classification. Math. Biosci. Eng. **16**(5), 3345–3367 (2019)
16. Zhang, X., Peng, F., Long, M.: Robust coverless image steganography based on DCT and LDA topic classification. IEEE Trans. Multimed. **20**(12), 3223–3238 (2018)
17. Wang, S., Sun, J., Phillips, P., Zhao, G., Zhang, Y.: Polarimetric synthetic aperture radar image segmentation by convolutional neural network using graphical processing units. Real-Time Image Process. **15**(3), 631–642 (2018)
18. Qi, L., Yu, J., Zhou, Z.: An invocation cost optimization method for web services in cloud environment. Sci. Program. (2017)
19. Zhou, Z., Jonathan, W., Sun, X.: Multiple distances-based coding: toward scalable feature matching for large-scale web image search. IEEE Trans. Big Data (2019)
20. Pan, L., Qin, J., Chen, H., Xiang, X., Li, C., Chen, R.: Image augmentation-based food recognition with convolutional neural networks. Comput. Mater. Contin. **59**(1), 297–313 (2019)
21. Pan, W., Qin, J., Xiang, X., Wu, Y., Tan, Y., Xiang, L.: A smart mobile diagnosis system for citrus diseases based on densely connected convolutional networks. IEEE Access. **7**, 87534–87542 (2019)
22. Xu, Y., Qi, L., Dou, W., Yu, J.: Privacy-preserving and scalable service recommendation based on SimHash in a distributed cloud environment. Complexity (2017)
23. Zhang, J., Jin, X., Sun, J., Wang, J., Sangaiah, A.K.: Spatial and semantic convolutional features for robust visual object tracking. Multimed. Tools Appl. (2018)
24. Qin, J., Li, H., Xiang, X., Tan, Y., Pan, W., Xiong, N.N.: An encrypted image retrieval method based on harris corner optimization and LSH in cloud computing. IEEE Access. **7**(1), 24626–24633 (2019)
25. Qi, L., Zhang, X., Dou, W., Ni, Q.: A distributed locality-sensitive hashing based approach for cloud service recommendation from multi-source data. IEEE J. Select. Areas Commun. **35**(11), 2616–2624 (2017)
26. Xiang, L., Shen, X., Qin, J., Hao, W.: Discrete multi-graph hashing for large-scale visual search. Neural Process. Lett. **49**(3), 1055–1069 (2019)
27. Yuan, C., Xia, Z., Sun, X.: Coverless image steganography based on SIFT and BOF. J. Internet Technol. **18**(2), 435–442 (2017)
28. Li, H., Qin, J., Xiang, X., Pan, L., Ma, W., XIONG, N.N.: An efficient image matching algorithm based on adaptive threshold and RANSAC. IEEE Access. **6**(1), 66963–66971 (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Yuanjing Luo received her B.S. in automation from Hainan Normal University, China, in 2018. She is currently pursuing her M.S. in Computer Technology at College of Computer Science and Information Technology, Central South University of Forestry and Technology, China. Her research focuses on deep learning and image processing.

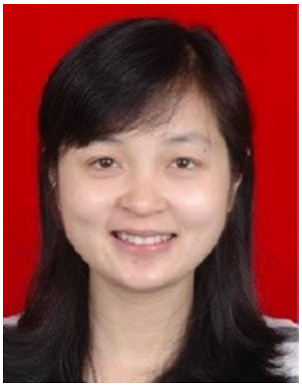
Jiaohua Qin received the B.S. in mathematics from the Hunan University of Science and Technology, China, in 1996, the M.S. degree in computer science and technology from the National University of Defense Technology, China, in 2001, and the Ph.D. degree in computing science from Hunan University, China, in 2009. She was a Visiting Professor with the University of Alabama, Tuscaloosa, AL, USA, from 2016 to 2017. She is currently a professor with the College of Computer Science and Information Technology, Central South University of Forestry and Technology, China. Her research interests include network and information security, machine learning and image processing.

Xuyu Xiang received his B.S. in mathematics from Hunan Normal University, China, in 1996, M.S. degree in computer science and technology from National University of Defense Technology, China, in 2003, and Ph.D. in computing science from Hunan University, China, in 2010. He is a professor with the College of Computer Science and Information Technology, Central South University of Forestry and Technology, China. His research interests include network and information security, image processing and machine learning.



Yun Tan received the M.S. and Ph.D. degrees both from Beijing University of Posts and Telecommunications, China, in 2004 and 2016, respectively. Now she is a lecturer in College of Computer Science and Information Technology, Central South University of Forestry and Technology. Her research interests mainly include image security, compressive sensing and signal processing.

Qiang Liu received his B.S. in network engineering from Hunan University of Technology, China, in 2017. He is currently pursuing his M.S. in Computer Technology at College of Computer Science and Information Technology, Central South University of Forestry and Technology, China. His research interests include machine learning and image processing.



Lingyun Xiang received the B.E. degree in computer science and technology and the Ph.D. degree in computer application from Hunan University, Hunan, China, in 2005 and 2011, respectively. She is currently a Lecturer with the School of Computer and Communication Engineering, Changsha University of Science and Technology. Her research interests include information security, steganography, steganalysis, machine learning, and pattern recognition.